

Trabajo de Fin de Grado

Integración de sistemas para el análisis de redes sociales utilizando procesamiento de lenguaje natural.

Constanza Polette León Baritussio

La Laguna, 29 de Marzo de 2017

D. **Dagoberto Castellanos Nieves**, con N.I.F. 79234766-L profesor Ayudante Doctor de Universidad adscrito al Departamento de Ingeniería Informática y de Sistemas de la Universidad de La Laguna, como tutor.

**C E R T I F I C A (N)**

Que la presente memoria titulada:

*“Estudio de similitudes semánticas en Redes Sociales.”*

ha sido realizada bajo su dirección por D. **Constanza Polette León Baritussio**, con N.I.F. X-6337312-F.

Y para que así conste, en cumplimiento de la legislación vigente y a los efectos oportunos ﬁrman la presente en La Laguna a 7 de Febrero de 2017.

Agradecimientos

XXX

XXX

XXX

XXX

Licencia

\* Si NO quiere permitir que se compartan las adaptaciones de tu obra y NO quieres permitir usos comerciales de tu obra indica:



© Esta obra está bajo una licencia de Creative Commons Reconocimiento-NoComercial-SinObraDerivada 4.0 Internacional.

\* Si quiere permitir que se compartan las adaptaciones de tu obra mientras se comparta de la misma manera y NO quieres permitir usos comerciales de tu obra indica:

C:\Users\Usuario\Desktop\memoriaTFG_1415_LaTeX\tfg\images\by-nc-sa_88x31.png

© Esta obra está bajo una licencia de Creative Commons Reconocimiento-NoComercial-CompartirIgual 4.0 Internacional.

\* Si quiere permitir que se compartan las adaptaciones de tu obra y NO quieres permitir usos comerciales de tu obra indica:

C:\Users\Usuario\Desktop\memoriaTFG_1415_LaTeX\tfg\images\by-nc_88x31.png

© Esta obra está bajo una licencia de Creative Commons Reconocimiento-NoComercial 4.0 Internacional.

\* Si NO quiere permitir que se compartan las adaptaciones de tu obra y quieres permitir usos comerciales de tu obra indica:

C:\Users\Usuario\Desktop\memoriaTFG_1415_LaTeX\tfg\images\by-nd_88x31.png

© Esta obra está bajo una licencia de Creative Commons Reconocimiento-SinObraDerivada 4.0 Internacional.

\* Si quiere permitir que se compartan las adaptaciones de tu obra mientras se comparta de la misma manera y quieres permitir usos comerciales de tu obra (licencia de Cultura Libre) indica:

C:\Users\Usuario\Desktop\memoriaTFG_1415_LaTeX\tfg\images\by-sa_88x31.png

© Esta obra está bajo una licencia de Creative Commons Reconocimiento-CompartirIgual 4.0 Internacional.

\* Si quiere permitir que se compartan las adaptaciones de tu obra y quieres permitir usos comerciales de tu obra (licencia de Cultura Libre) indica:

C:\Users\Usuario\Desktop\memoriaTFG_1415_LaTeX\tfg\images\by_88x31.png

© Esta obra está bajo una licencia de Creative Commons Reconocimiento 4.0 Internacional.

**Resumen**

*El objetivo de este trabajo ha sido, desarrollar una integración de sistemas, en concreto un sistema de alerta y un sistema de análisis de sentimiento.*

*La competencia [E6], que figura en la guía docente, indica que en la memoria del trabajo se ha de incluir: antecedentes, problemática o estado del arte, objetivos, fases y desarrollo del proyecto, conclusiones, y líneas futuras.*

*El documento de memoria debe tener un máximo de 50 páginas.*

*No se deben dejar páginas en blanco al comenzar un capítulo, ya que el documento no está pensado para sea impreso sino visionado con un lector de PDFs.*

*También es recomendable márgenes pequeños ya que, al firmar digitalmente por la Sede, se coloca un marco alrededor del texto original.*

*El tipo de letra base ha de ser de 14ptos.*

**Palabras clave:** Red Social, similitud semántica.

**Abstract**

*The objective of this work has been,the extraction of data from the social network Twitter,it is a perfoming tool capable of analytics on the semantic similaritis between the concepts used on social networks.*

***Keywords:*** *Social network, semantic similarity.*

Índice General

[Capitulo 1. Introducción 12](#_Toc478404521)

[1.1 Internet 12](#_Toc478404522)

[1.2 Web 2.0 12](#_Toc478404523)

[1.3 Redes Colaborativas 14](#_Toc478404524)

[1.4 Redes Sociales 15](#_Toc478404525)

[Capítulo 2. Sistemas de alerta y sistemas de análisis de sentimientos. 17](#_Toc478404526)

[Capítulo 3. 18](#_Toc478404527)

[Integración de los sistemas de alerta y sistemas de análisis de sentimientos propuestos. 18](#_Toc478404528)

[1.5 Primer apartado de este capítulo 18](#_Toc478404529)

[1.6 Segundo apartado de este capítulo 18](#_Toc478404530)

[1.7 Tercer apartado de este capítulo 18](#_Toc478404531)

[Capítulo 4. Recomendaciones o conclusiones. 19](#_Toc478404532)

[Conclusiones y líneas futuras 20](#_Toc478404533)

[Summary and Conclusions 21](#_Toc478404534)

[1.8 First Section 21](#_Toc478404535)

[Presupuesto 22](#_Toc478404536)

[1.9 Sección Uno 22](#_Toc478404537)

[Apéndice A. Título del Apéndice 1 23](#_Toc478404538)

[A.1. Algoritmo XXX 23](#_Toc478404539)

[A.2. Algoritmo YYY 23](#_Toc478404540)

[Apéndice B. Título del Apéndice 2 25](#_Toc478404541)

[B.1. Otro apendice: Seccion 1 25](#_Toc478404542)

[B.2. Otro apendice: Seccion 2 25](#_Toc478404543)

[Bibliografía 26](#_Toc478404544)

Índice de figuras

[Figura 1.1. Primera figura (el formato es **Leyenda**). 3](#_Toc418604435)

[Figura 1.2. Panel de Estilos de la plantilla TFG. 3](#_Toc418604436)

Índice de tablas

[Tabla 1.1. Esta es la primera tabla con estilo **Tabla 1** (usar estilo **Celda de tabla** y luego dar formato). El encabezado tiene estilo **Encabezado de tabla**. 3](#_Toc418604518)

[Tabla 1.2. Segunda tabla con estilo **Tabla 2** (ídem). 3](#_Toc418604519)

[Tabla 7.1. Tabla resumen de los Tipos. 3](#_Toc418604520)

## Capitulo 1. Introducción

### Internet

El nombre Internet proviene de las palabras en inglés Interconnected Networks, que significa "redes interconectadas". Sus orígenes se remontan a 1969, cuando se estableció la primera conexión entre ordenadores, conocida como ARPANET.

Internet, es una arquitectura de sistema que ha revolucionado las comunicaciones y los métodos de comercio al permitir que varias redes de computadoras estén interconectadas. Proporciona un mecanismo potente y accesible por cualquier individuo, ya que, ofrece una amplía información que puede ser utilizada para cualquier propósito en general. La difusión de textos digitalizados, imágenes y grabaciones de audio y vídeo, ha dado lugar a una explosión de información [1].

El internet, se ha convertido en uno de los medios más importantes desde finales del siglo 20 y principios del 21 en la sociedad, ya que, trajo consigo una nueva forma de comunicación entre las personas y una mejora en la transmisión y acceso de la información eliminando así las barreras de espacio/tiempo, surgiendo así la World WideWeb (Tim Berners-Lee y Robert Cailliau 1990) o también conocida como Web 1.0.

### Web 2.0

Después de la llegada de la World Wide Web en los 90 y la Web2.0 (Tim O´Reilly y Dale Dougherty en 2004), para referirse a una segunda generación en la historia de la Web. La Web 2.0 se puede entender como –“todas aquellas utilidades y servicios de Internet que se sustentan en una base de datos, la cual puede ser modificada por los usuarios del servicios, ya sea en su contenido (añadiendo, modificando o borrando información o asociando datos a la información existente), pues bien en la forma de presentarlos, o en contenido y forma simultáneamente”- (Ribes, 2007). Se puede afirmar que el término “Web 2.0” representa la evolución de las aplicaciones tradicionales hacia aplicaciones web enfocadas en el usuario final. No se trata de una nueva tecnología sino de una actitud de colaboración y participación de las personas para proporcionar mejores datos, nuevos servicios y aplicaciones on-line [2]. En donde los usuarios pueden colaborar e interactuar entre sí [3], creando contenidos dentro de una comunidad virtual y conectándose con otros individuos.

Aunque todos los sitios de la Web están interconectados entre sí, con el fin de organizar los recursos digitales (Cobo y Pardo 2007)[4],proponen una taxonomía compuesta por cuatro líneas fundamentales que ayudará al usuario a seleccionar aquellas aplicaciones que cubran de una mejor manera sus necesidades e intereses. La estructura propuesta se denomina “los cuatro pilares de la Web 2.0”, que son:

* **Redes sociales:** describe todas aquellas herramientas diseñadas para la creación de espacios que promuevan o faciliten la formación de comunidades e instancias de intercambio social.
* **Contenidos:** hace referencia a aquellas herramientas que favorecen la escritura en línea, así como su distribución e intercambio.
* **Organización social e inteligente de la información:** herramientas y recursos para etiquetar, sindicar e indexar, que faciliten el orden y almacenamiento de la información, así como otros recursos disponibles en la Red.
* **Aplicaciones y servicios (mashups):** dentro de esta clasificación se incluye numerosas herramientas, software, plataformas en línea y un hibrido de recursos creados para ofrecer servicios de valor al usuario final.

### Redes Sociales

Como se describía en la sección anterior, la Web 2.0 está compuesta de comunidades de usuarios y una gama especial de servicios, como por ejemplo, las redes sociales como Linkedin, Instagram, los blogs como Blogger, Wordpress y las wikis como Wikipedia.

Las redes sociales, se pueden definir como una estructura compuesta por un grupo de individuos u organizaciones y que están relacionados entre sí. Dichas relaciones pueden ser de tipo: laborales, amistad, familiar, entre otros. La creación de las redes sociales supone una nueva forma de comunicación. Estas herramientas ofrecen un espacio virtual para que, miles de personas pueden compartir contenidos multimedia con intereses similares (Granovetter, 1973) [5]. Se consideran puntos de encuentro, en los que es posible acceder a la información, compartir impresiones, consultar archivos y recursos disponibles a tiempo real (Nass De Ledo, 2011)[6]. Redes sociales como Facebook, Twitter, Tinder, son estructuras compuestas por una infinidad de usuarios relacionados entre sí, independientemente que se conozcan o no. En estas redes sociales, al tener un gran número de usuarios, se manejan grandes volúmenes de datos, ya que, los individuos comparten, buscan y difunden información a través de texto, imágenes o audio. Gracias a las redes sociales, las personas pueden expresar opiniones digitales en foros, debates y hacer que estas lleguen a un público muy amplio y diverso

En 2002, comenzaron a aparecer los primeros sitios web, promocionando las redes de círculos de amigos en línea. La popularidad de estos sitios creció rápidamente y compañías como por ejemplo, Google lanzó Orkut en 2004.

En la actualidad existen diversas redes sociales que se clasifican según la tipología:

* **Horizontales:** provee herramientas para la interrelación general: Facebook.
* **Redes sociales humanas:** Fomentar las interacciones entre distintas personas según gustos, actividades, por ejemplo Dopplr.
* **Redes sociales de contenido:** El relacionamiento de los usuarios se establece n función del contenido que se publica en la red, por ejemplo Flickr.
* **Redes sociales sedentarias:** Son redes que dependen de los eventos, y contenidos publicados, como por ejemplo Blogger.
* **Redes sociales nómadas:** Depende de la localización geográfica del usuario. Por ejemplo: Tinder.
* **Verticales**
  + **Tipo usuario:** Dirigido a un grupo específico, personas que comparten intereses en productos o servicios. Por ejemplo Parobook.
  + **Tipo de actividad:** Son redes que promueven opiniones sobre productos o servicios, por ejemplo Twitter.
  + **Profesionales:** Redes especializadas en establecer relaciones laborales, por ejemplo Linkedin.
  + **Ocio:** Se trata de conectar a los individuos que tengan gustos similares en algún área como la música o el deporte. Por ejemplo, Dogster.
  + **Mixtas:** Consiste en unificar dos o más redes verticales, descritas anteriormente, por ejemplo Unience.

Uno de los objetivos de la creación de estas redes, es crear un sitio de interacción virtual en donde miles de personas de todo el mundo puedan compartir, comunicarse y cooperar entre sí.

### Redes Colaborativas

Una red colaborativa se compone de un conjunto de personas que hacen alguna aportación intelectual a un proyecto o investigación con un objetivo en común con el grupo que colabora.

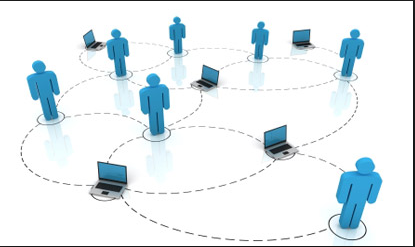


Figura 3: Red Colaborativa

Existen diferentes modelos de redes colaborativas para la innovación, según los objetivos que se quieran cumplir, denominados:

* **Modelo Orquesta:** las redes colaborativas que siguen este modelo, describen un grupo de compañías que obtienen en conjunto una oportunidad de innovar algo, y el liderazgo es asumido por una compañía líder. Tiende a ser bien organizado y coordinado con una alta inversión en infraestructuras.
* **Bazar Creativo:** en este modelo, las compañías compran ideas, productos y tecnologías innovadoras en lugar de crearlas. Para lanzarlo al mercado rápidamente y así, obtener beneficios y remuneraciones de manera inmediata. El modelo de liderazgo sigue un modelo de firma dominante.
* **Central de aglomeración:** este modelo de innovación, las contribuciones se hacen de manera individual. El desarrollo de las ideas innovadoras, es muy rápido gracias a la colaboración de los múltiplos miembros de la comunidad. No existe una responsabilidad de liderazgo, ya que, todos los miembros del equipo participan.
* **Estación de MOD:** las redes colaborativas que optan por este modelo, poseen dos características importantes, la primera es una involucración total en mejorar la innovación que ya existe, ya sea, un producto, servicio o proceso. Y la segunda, es que, la innovación suceda en una comunidad en la cual los beneficiarios sean los miembros de la red y no para la firma dominante. Este modelo sigue un liderazgo descentralizado y sobre una plataforma ya previamente establecida, esto permite la generación de nuevos productos a partir de la misma.

## Capítulo 2. Sistemas de alerta y sistemas de análisis de sentimientos.

## Capítulo 3.

## Integración de los sistemas de alerta y sistemas de análisis de sentimientos propuestos.

En el capítulo anterior se ha introducido conceptos como Internet, Web 2.0, Redes colaborativas y Redes Sociales, ahora se procederá a la extracción de la información utilizando el lenguja R

### Primer apartado de este capítulo

### Segundo apartado de este capítulo

### Tercer apartado de este capítulo

## Capítulo 4. Recomendaciones o conclusiones.

Los capítulos intermedios servirán para cubrir los siguientes aspectos: antecedentes, problemática o estado del arte, objetivos, fases y desarrollo del proyecto.

En el capítulo 1 se describió bla, bla, bla.....

## Conclusiones y líneas futuras

Este capítulo es obligatorio. Toda memoria de Trabajo de Fin de Grado debe incluir unas conclusiones y unas líneas de trabajo futuro.

## Summary and Conclusions

This chapter is compulsory. The memory should include an extended summary and conclusions in English.

### First Section

## Presupuesto

Este capítulo es obligatorio. Toda memoria de Trabajo de Fin de Grado debe incluir un presupuesto.

### Sección Uno

|  |  |
| --- | --- |
| Tipos | Descripción |
| AAA | BBB |
| CCC | DDD |
| EEE | FFF |
| GGG | HHH |

|  |  |
| --- | --- |
| Tipos | Descripción |
| AAA | BBB |
| CCC | DDD |
| EEE | FFF |
| GGG | HHH |

Tabla 7.1. Tabla resumen de los Tipos.

## Apéndice A. Título del Apéndice 1

### A.1. Algoritmo XXX

/\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*

\*

\* Fichero .h

\*

\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*

\*

\* AUTORES

\*

\*

\* FECHA

\*

\*

\* DESCRIPCION

\*

\*

\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*/

### A.2. Algoritmo YYY

/\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*

\*

\* Fichero .h

\*

\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*

\*

\* AUTORES

\*

\*

\* FECHA

\*

\*

\* DESCRIPCION

\*

\*

\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*/

## Apéndice B. Título del Apéndice 2

### B.1. Otro apendice: Seccion 1

Texto

### B.2. Otro apendice: Seccion 2

Texto

## Bibliografía

1. "Internet." Britannica Academic, Encyclopædia Britannica, 15 Mar. 2016.
2. Flores Cueto, J. J., Morán Corzo, J. J., & Rodríguez Vila, J. J. (2009). Las redes sociales. *Universidad de San Martín de Porres*, 1-15.
3. Prato, Laura Beatriz: *“ Aplicaciones Web 2.0: redes sociales ”* [987-1727-13-5; 1-4492-7744-6] p.1. (2010).
4. COBO, C. & PARDO, H. Planeta Web 2.0. Inteligencia colectiva o medios fast food (Grup de Recerca d'Interaccions Digitals, Universitat de Vic. Flacso México.). Barcelona/México DF. (2007).
5. Prato, Laura Beatriz: *“ Aplicaciones Web 2.0: redes sociales ”* [987-1727-13-5; 1-4492-7744-6] p.16. (2010).
6. Freire, E. S., Ordóñez, F. S., & Martínez, M. P: “Redes sociales, un mecanismo de impulso del emprendimiento turístico en la provincia Tungurahua, Ecuador/Social networks to promote the tourism project in The Tungurahua Province, Ecuador”. *Retos Turísticos*, *15*(3). (2017).